Erschienen in: Cognitive Linguistics Bd. 19 (2008) H. 3, S. 349–355

Introduction

ARNE ZESCHEL*

One of the central tenets of Cognitive Linguistics is its fundamentally usage-based orientation: language is seen as an inventory of dynamic symbolic conventions (constructions) whose organisation is constantly updated by (and hence adapting to) language use (Langacker 2000). Such usage-based, "emergentist" views of language are also found in recent work outside Cognitive Linguistics in the narrower sense: for instance, there is experimental evidence from various sources that shared symbolic communication systems can indeed emerge (on the interpersonal level) and be learned (on the individual level) in a data-driven, self-organising manner that does not require substantial language-specific stipulations (be it in humans or machines).[1] This is not to deny that many aspects of the usage-based language model are still underspecified and have the status of assumptions rather than established facts. However, there is currently a commendable trend within Cognitive Linguistics to put its programmatic appeal to the usage-based hypothesis to the test: more and more studies set out to evaluate specific predictions of the approach in different domains against appropriate experimental and/or corpus data, thereby contributing to a successive refinement of the overall model and helping to put it on a sound empirical footing (cf. Tummers et al. 2005 as well as the papers in Gries and Stefanowitsch 2006 and González-Márquez et al. 2007 for recent overviews and applications).

* Author's e-mail address: zeschel@uni-bremen.de. I am grateful to Kerstin Fischer, Anatol Stefanowitsch and Felix Bildhauer for comments.

1. For the spontaneous emergence of novel symbolic communication systems among humans, cf. Galantucci (2005); for the emergence of shared linguistic communication systems (construction grammars) among cognitive robots, cf. Steels (2005); for overviews of the usage-based approach to child language acquisition, cf. Tomasello (2003) and Goldberg (2006); for unsupervised machine learning of a Langacker-style natural language construction grammar, cf. Solan et al. (2005).

The papers in this special issue (which has grown out of a theme session on "Constructions in Language Processing" held at the 2nd International Conference of the German Cognitive Linguistics Association in Munich in October 2006) all represent this line of research, with a focus on constructionist perspectives on (human) language processing and its relationship to the linguistic representations that speakers extract from their experience.

The issue opens with a study of island effects in English clausal complement constructions by *Ben Ambridge and Adele Goldberg* ("The island status of clausal complements: evidence in favor of an information structure explanation"). The authors compare the classical subjacency account of constraints on filler-gap relations (Chomsky 1973) with an item-based analogical approach (in which acceptability is a function of semantic distance to a stored prototype) and their own proposal, in which ease of extraction depends on the target's degree of "backgroundedness" in discourse (a principle which they refer to as "BCI: backgrounded constituents are islands"). Ambridge and Goldberg substantiate their hypothesis with the results of two questionnaire studies, suggesting that the effects investigated are best interpreted as a pragmatic anomaly reflecting the fact that a constituent cannot be at the same time backgrounded and focused. The authors conclude that the possibility of combining two constructions in production is influenced by the information-structural properties of the constructions involved (among other things).

Unbounded dependency constructions in English are also the topic of the second study, "Questions with long-distance dependencies: A usage-based perspective" by *Ewa Dąbrowska.* In contrast to Ambridge and Goldberg, Dąbrowska is concerned with how the acceptability of different types of WH-questions with long-distance dependencies can be predicted from their similarity to an assumed prototype rather than from general semantic/pragmatic principles: departing from the observation that naturally occurring instances of this construction tend to be highly stereotypical, she suggests that they are not derived by abstract rules but by modifying (or, in comprehension: by comparing a given target to) a stored low-level schema of the format *WH do you think/say S-GAP?* Dąbrowska presents evidence for the predicted prototypicality effects from an acceptability judgment experiment and points to possible interpretations of the obtained results in terms of both strongly item-based/analogical models and a hybrid architecture that also represents abstract schemas alongside specific exemplars.

Similar to the first two contributions, the third and fourth paper in the volume both deal with the same linguistic phenomenon, but with a different focus and with different aims. In my own contribution ("Lexical

chunking effects in syntactic processing''), I report an experiment on syntactic ambiguity resolution that seeks to probe the psychological reality and processing relevance of partially schematic prefabs (i.e., the kinds of low-level schemas that speakers are assumed to store in usage-based Construction Grammar). The results of the experiment indicate that global complementation preferences applying to a given verb ''at large'' (i.e., considering its entire usage spectrum) may be overridden by conflicting evidence for specific syntagmatic chunks in which this verb occurs. These results are interpreted as support for the usage-based view that such structures may have independent memory storage even when they are fully predictable, and that such representations are furthermore privileged over more abstract (i.e., lexically unfilled) constructions in language processing.

Dealing with the same phenomenon (i.e., garden path effects resulting from a specific type of local syntactic ambiguity in English), *Daniel Wiechmann*'s paper ''Initial parsing decisions and lexical bias: Corpus evidence from local NP/S-ambiguities'' has a more methodological focus. The author presents a corpus-linguistic approach to assessing verbal complementation preferences in terms of collostruction strength using the method of Distinctive Collexeme Analysis (DCA; Gries and Stefanowitsch 2004). Using a balanced corpus, both verb-general and (verb-)sense-specific associations with different complementation patterns are computed for 20 verbs and related to on-line measures of processing difficulty from an earlier reading experiment with these verbs (Hare et al. 2003). The results confirm the hypothesis that sense-specific associations (as determined by the DCA) are a better predictor of processing preferences/difficulties than form-based associations. Moreover, the author suggests that the observed correlation between the corpus-derived predictions and Hare et al.'s experimental findings indicates that collostruction strength is a valid approximation of constructional association strength on the psychological plane.

Holger Diessel's study ''Iconicity of sequence: A corpus-based analysis of the positioning of temporal adverbial clauses in English'' is devoted to aspects of production again. The author discusses a range of factors that influence speakers' choice of the positioning of adverbial clauses relative to the matrix clause in different languages, with special attention to one of these motivations, iconicity of sequence (i.e., the iconic encoding of prior events in preposed clauses and posterior events in postposed clauses). Diessel's study reveals that the ordering of temporal adverbial clauses in English is significantly affected by iconicity of sequence, which is viewed as a processing principle geared at avoiding structures that are difficult to plan and to interpret. In a second step, the author uses logistic

regression analysis to integrate the observed effect into a more comprehensive model of processing constraints on clause order in complex sentences which also includes factors such as clause length, syntactic complexity and pragmatic import. The resulting picture is a model in which speakers seek to balance multiple constraints on their constructional encoding options in order to minimise overall processing load.

Though concerned with yet a different aspect of language processing, *Martin Hilpert*'s study "New evidence against the modularity of grammar: Constructions, collocations and speech perception" is again interested in the psychological status and processing relevance of entrenched exemplars of a given construction. However, the overall thrust of Hilpert's argument is different from that of other papers in the issue which are concerned with item-based effects in language processing: by showing that the phonemic categorisation of a synthesised ambiguous sound (located somewhere on a continuum between two phonemes) can be biased in either direction by embedding it in an appropriate collocational "carrier phrase", the study documents syntactic top-down effects on word recognition that are difficult to reconcile with strictly serial-modular theories of language processing. Hilpert provides evidence that the observed effect applies immediately (i.e., at the level of auditory input processing), which implies that it cannot be explained by appealing to "late feedback" between modules. Instead, the author argues that frequent word combinations have psychological reality as independent units of linguistic knowledge, and that lexical and syntactic aspects of language processing are not plausibly attributed to separate (i.e., "informationally encapsulated") mental modules.

Like Hilpert's study, the final contribution addresses a famous tenet of linguistic theories that are decidedly *non*-emergentist: in "Negative entrenchment: A usage-based approach to negative evidence", *Anatol Stefanowitsch* presents a new perspective on the so-called "no negative evidence" problem that figures prominently in nativist accounts of language acquisition. The author contrasts different strategies for overcoming the problem that have been proposed in the literature and then presents a new approach that builds on the notion of "negative entrenchment": if speakers keep track of how often a particular simplex element or feature *occurs* in the input, Stefanowitsch argues, such information could be used to form subconscious expectations as to how often it should *co-occur* with other elements or features in the language if there were nothing in the grammar to prevent this. Learners could thus distinguish absences in the input that are statistically significant from those that are merely accidental, with continued non-occurrence of statistically expected combinations resulting in their growing "negative entrenchment". The

author backs up his proposal with the results of a pilot study which suggests that corpus-derived scores of negative entrenchment are a better predictor of experimental (un)acceptability judgments than corpus-derived measures of constructional pre-emption (i.e., one of the other mechanisms discussed in the literature that are assumed to compensate for the lack of explicit negative evidence).

In sum, the papers collected in this special issue demonstrate many interesting prospects of combining a usage-based approach to grammar with suitable empirical methodologies: the contributions fill empirical and methodological gaps on the constructionist research agenda (Wiechmann; Diessel), they put important assumptions of the hypothesised model to the test or extend it in novel ways (Zeschel; Hilpert; Stefanowitsch), they reframe classical issues in grammatical theory from a usage-based perspective (Ambridge and Goldberg; Dąbrowska; Stefanowitsch), and they challenge more general claims about the properties of language and cognition that rest in part on questionable arguments from theoretical linguistics (Hilpert). At the same time, there are a number of important issues on which not all contributors might agree (such as the scope and explanatory status of item-based approaches to language processing and representation; cf. Abbot-Smith and Tomasello 2006). However, this should only encourage further empirical investigation of these issues, and future research can of course only benefit from the fact that relevant differences are clearly articulated rather than glossed over.

That said, readers may wonder how it is that one particular strand of this research is not featured in this special issue at all – i.e., usage-based work in computational linguistics. Clearly, statistical approaches to natural language processing share important assumptions of usage-based theories of language, and particular models might thus provide a useful empirical touchstone for hand-crafted reconstructions of e.g., construction learning processes (cf. Bod, in press). Moreover, moving beyond the purely statistical aspects of language and language processing, the transition from traditional computational modelling to experiments with embodied robotic agents that learn to associate linguistic signs with aspects of their sensory-motor experience (e.g., Dominey and Boucher 2005; Steels and Kaplan 2002; Sugita and Tani 2005) provides a wealth of further interesting possibilities for investigating some of the very key concerns of Cognitive Linguistics from a new perspective (cf. also Zlatev and Balkenius 2001). However, it is beyond the scope of this special issue to map out points of contact between these two research communities. For the moment, then, suffice it to acknowledge that usage-based approaches to language are gaining more and more currency also in neighbouring disciplines, and that the increasing integration of appropriate

methodologies from linguistics, cognitive psychology and computer science promises many interesting perspectives for future research on the cognitive instantiation of language.

Universität Bremen, Germany

References

Abbot-Smith, Kirsten, and Michael Tomasello
2006 Exemplar-learning and schematization in a usage-based account of syntactic acquisition. *The Linguistic Review* 23, 275–290.

Bod, Rens
in press Constructions at work or at rest? *Cognitive Linguistics.*

Chomsky, Noam
1973 Conditions on Transformations. In: Anderson, Stephen R., and Paul Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt, Rinehart & Winston, 232–286.

Dominey, Peter F. and Jean-David Boucher
2005 Developmental stages of perception and language acquisition in a perceptually grounded robot. *Cognitive Systems Research* 6, 243–259.

Galantucci, Bruno
2005 An experimental study of the emergence of human communication systems. *Cognitive Science* 29, 737–767.

Goldberg, Adele E.
2006 *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.

González-Márquez, Mónica, Irene Mittelberg, Seana Coulson and Michael J. Spivey (eds.)
2007 *Methods in Cognitive Linguistics.* Amsterdam: John Benjamins.

Gries, Stefan Th. and Anatol Stefanowitsch
2004 Extending collostructional analysis: a corpus-based perspective on 'alternations'. *International Journal of Corpus Linguistics* 9(1), 97–129.

Gries, Stefan Th., and Anatol Stefanowitsch (eds.)
2006 *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*. Berlin: Mouton de Gruyter.

Hare, Mary L., Ken McRae, and Jeffrey L. Elman
2003 Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory and Language* 48(2), 281–303.

Langacker, Ronald W.
2000 A dynamic usage-based model. In: Barlow, Michael and Suzanne Kemmer (eds.), *Usage-Based Models of Language*. Stanford: CSLI, 1–63.

Solan, Zach, David Horn, Eytan Ruppin, and Shimon Edelman
2005 Unsupervised learning of natural languages. *Proc. Natl. Acad. Sci.* 102, 11629–11634.

Steels, Luc
2005 The emergence and evolution of linguistic structure: from lexical to grammatical communication systems. *Connection Science* 17(3/4), 213–230.

Steels, Luc and Frederic Kaplan
2002 Bootstrapping grounded word semantics. In Briscoe, Ted (ed.), *Linguistic evolution through language acquisition: formal and computational models*. Cambridge: Cambridge University Press, 53–73.

Sugita,Yuuya and Jun Tani
2005 Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adaptive Behavior* 13(1), 33–52.
Tummers, José, Kris Heylen, and Dirk Geeraerts
2005 Usage-based approaches in Cognitive Linguistics: A technical state of the art. *Corpus Linguistics and Linguistic Theory* 1, 225–261.
Tomasello, Michael
2003 *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.
Zlatev, Jordan, and Christian Balkenius
2001 Introduction: Why Epigenetic Robotics? In: Balkenius, Christian, Jordan Zlatev, Hideki Kozima, Kerstin Dautenhahn, and Cynthia Breazeal (eds.), *Proceedings of the First International Workshop on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*. Lund University Cognitive Studies, 85 1–4.